An Efficient Salient Region Detection on Underwater Image with Grabcut Algorithm

Prasanna .S, Durgadevi .S, Senthilkumar .S

Abstract*: Saliency detection aims at automatically estimating visually salient object regions in an image, saliency segmentation and foreground extraction are two important applications of this. However, it is a challenge for underwater images to estimate salient regions by saliency detection methods because of the low-contrast and poor quality. In this paper, we address this problem by combining the detected object regions rather than the whole image, where Sobel edge detector and Active contour is used for proposing candidate regions. We extensively evaluated our method on underwater images, and experimental results show that the performances of saliency detection and segmentation are improved. These saliency segmentation masks are further used to extract the foreground objects of an image.

***Index Terms*: Saliency segmentation, foreground extraction, fish localization, underwater image.**

I. INTRODUCTION

Images are considered as one of the most important medium of conveying information. An image is an array, or a matrix of square pixels (picture elements) arranged in columns and rows. Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. The algorithms for image processing, developed over many years for land based applications, often make simplifying assumptions about the image data that are almost certainly not valid underwater. Image processing is a process where input image is processed to get output also as an image or attributes of the image. Main aim of all image processing techniques is to recognize the image or object under consideration easier visually. Segmentation of images holds a crucial position in the field of image processing. In medical imaging, segmentation is important for feature extraction, image measurements and image display. Image segmentation can be defined as the partition or segmentation of a digital image into similar regions with a main aim to simplify the image under consideration into something that is more meaningful and easier to analyse visually. Image segmentation methods can be classified as thresholding, region based, supervised and unsupervised techniques.

Revised Version Manuscript Received on June 05, 2017.

Prasanna.S, PG Scholar, ECE Department, E.G.S Pillay Engineering College, Nagappattinam, India. E-mail: dspras2076@gmail.com

Durgadevi. S, Assistant professor, E.G.S Pillay Engineering college, Nagappattinam, India, E-mail: durgadevi.me@gmail.com

Senthilkumar. S, Assistant professor, E.G.S Pillay Engineering college, Nagappattinam, India, E-mail: senthilkumar@egspec.org

In recent years, underwater image processing has drawn wide attentions. On one hand, with the developments of underwater imaging system, it becomes difficult and time-consuming to obtain the useful information from massive image data. On the other hand, a vast amount of underwater image data takes up a lot of storage and transmission resources. In order to efficiently solve these problems, we keep our eyes on foreground extraction of underwater images by saliency detection and segmentation in this paper.

As the name suggests, salient objects are the most visually noticeable foreground objects in the scene, while saliency detection means highlighting visually salient objects in an image. This fundamental task has drawn extensive attentions in recent years, which can be used to guide segmentation and extraction of interested regions of an image. In computer vision and image/video processing areas, saliency detection has a wide range of applications, such as image retrieval [5], content-aware image resizing [6], adaptive compression of image [7], etc. In addition, the video/image processing is under real-time with high accuracy in some applications, such as underwater image retrieval [5], transferring to the specific video frame and so on.

In order to locate the interested regions of an image, a foreground extraction of underwater image based salient objects detection is proposed. It is helpful not only to obtain useful information quickly, but also to manage storage and transmission resources efficiently. There are a large number of saliency detection approaches [1, 3, 8, 9, 10], most of which focus on design low-level cues and priors. Actually, these low-level features cannot help salient foreground objects stand out from confusing background in the low-contrast images. So the conventional saliency detection methods work not so well for underwater images. It is also inappropriate to apply the conventional methods for underwater image directly.

II. PROPOSED ALGORITHM

A. Underwater Images:

Water has a st rong back scattering characteristic and contains many suspended particulate impurities. These factors can influence the light translation in the water and cause the digital image background to be a non-uniform light environment. The conventional segmentation methods are not effective for the underwater image because they are sensitive to the noise and non uniform gray levels which exist in the image. With the theory and method of image processing developing, scholars find the mainly reason of the traditional image segmentation method can easily be affected by noise is that the image's useful information and noise are always overlapped in frequency band. Fish, which it is tempting to consider as objects, as in object recognition, are flexible, textured, and present themselves with a very wide range of poses and apparent sizes. They can appear anywhere in the field of view. We believe that existing image processing tools are far from ideal when applied to underwater images and that research on improving tool performance on underwater images is important.

Underwater image processing is a big challenge as depth of water increases the image quality decreases. Underwater images are more dull and dusty in look and we won't be able to identify the pictures clearly. Moreover most of the organisms in under water are not identified yet as there are millions of species in it. Underwater images suffer from a greater noise than normal images. New filters are being applied to remove this noise. Image noise is random variation of brightness or colour information in images and is usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse. This segmentation is done by Contrast based algorithm. The image is contrasted and a good quality image is obtained after the morphological image processing. The feature is extracted from the image using the Grabcut algorithm. The organism in underwater can be identified using this algorithm and foreground extraction is done by mapping the identified organism with the input image.

B. Detection and Segmentation of salient object:

In this paper, we focus on saliency segmentation and foreground extraction based on saliency detection. Saliency detection means to identify the most visually noticeable foreground objects in the scene quickly and accurately. As mentioned before, the conventional saliency detection methods [1, 3, 8, 9, 10] can achieve good results only for images with high quality. So a new efficient localization-based method is proposed for salient objects segmentation and foreground extraction (see Figure 5).

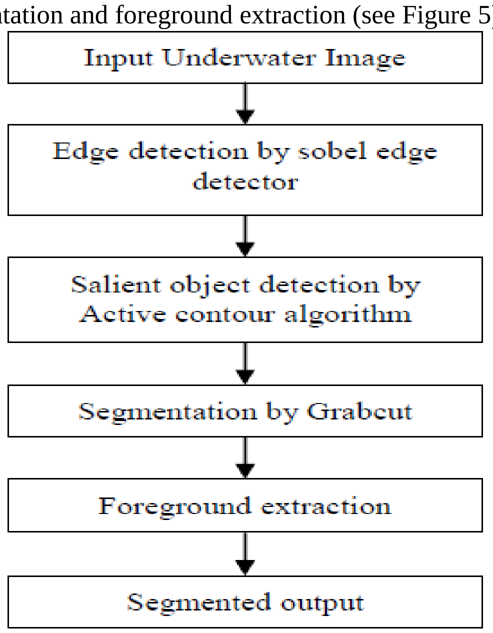

Figure 1: Process of detection and segmentation of a salient object

The main idea of this method is as follows:

First, we exploit the Active contour [13] to identify the fish in the detected salient regions of underwater images. The Sobel Edge detector [2] can be used for proposing those candidate regions, and detect the edges of the salient object. We segment the whole image into different parts based on the detected salient regions, and then we calculate the saliency maps, respectively.

Second, we run GrabCut [4] iteratively initialized by saliency maps to improve the saliency segmentation results and detect the exact region of salient object.

III. METHODOLOGY

A. The Input Underwater Image:

Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. The input is a two dimensional picture of an underwater image. The given input image contains more noise as it is underwater image. The sand or lack of light causes the picture to be unclear and difficult to identify the objects in the image. The Underwater image to be classified is selected as an input image from the corresponding file whiles it asking for an input.

B. Edge Detection:

In image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a features vector). This process is called feature selection. Feature extraction involves reducing the amount of resources required to describe a large set of data. Here, the desired features that is the edges are detected by using sobel operator.

Edge detection is a very important area in the field of computer vision. Edges define the boundaries between regions in an image, which helps with segmentation and object recognition. Edge detection is a fundamental of low-level image processing and good edges are necessary for higher level processing. The quality of edge detection is highly dependent on lighting conditions, the presence of objects of similar intensities, density of edges in the scene, and noise. Since different edge detectors work better under different conditions, it would be ideal to have an algorithm that makes use of multiple edge detectors, applying each one when the scene conditions are most ideal for its method of detection.

The basic idea of edge detection is as follows: First, use edge enhancement operator to highlight the local edge of the image. Then, define the pixel "edge strength" and set the threshold to extract the edge point set. However, because of the noise and the blurring image, the edge detected may not be continuous. So, edge detection includes two contents. First one is using edge operator to extract the edge point set. Second one is removing some of the edge points from the edge point set, filling it with some another and linking the obtained edge point set into lines.

The four steps of edge detection are:

Step 1: Smoothing suppress noise as possible, without destroying the true edges.

Step 2: Enhancement: apply a filter to enhance the quality of the edges in the image.

Step 3: Detection: determine which edge pixels should be discarded as noise and which should be retained.

Step 4: Localization: determine the exact location of an edge.

Before the edges are detected the input images are converted into gray scale image. Edge detection is the process of localizing pixel intensity transitions. The edge detection has been used by object recognition, target tracking, segmentation, and etc. There mainly exist several edge detection methods: Sobel, Prewitt, Roberts and Canny. In this paper, Sobel which is an edge detection method is considered. The Sobel edge detector uses two masks, one vertical and one horizontal. These masks are generally used 3×3 matrices.

Sobel operator is a kind of orthogonal gradient operator. Gradient corresponds to first derivative, and gradient operator is a derivative operator. Here, the image is convolved with only two kernels, one estimating the gradient in the x-direction, G_x, the other the gradient in the y-direction, G_y. The absolute gradient magnitude is then given by

$$|G| = \sqrt{G_{x^2} + G_{y^2}} \quad (1)$$

$$|G| = |G_x| + |G_y| \quad (2)$$

In many implementations, the gradient magnitude is the only output of a gradient edge detector. After having calculated the magnitude of the 1st derivative, we now have to identify those pixels corresponding to an edge. The easiest way is to threshold the gradient image, assuming that all pixels having a local gradient above the threshold must represent an edge. For a continuous function f (x, y), in the position (x, y), its gradient can be expressed as a vector.

$$\nabla f(x,y) = [G_x G_y] = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \quad (3)$$

The magnitude and direction angle of the vector are

$$mag(\nabla f) = |\nabla f_{(2)}| = [G_x{}^2 G_y{}^2]^{1/2} \quad (4)$$

$$\phi(x,y) = \arctan\left(\frac{G_x}{G_y}\right) \quad (5)$$

The partial derivatives of the formulas above need to be calculated for each pixel location. In practice, we often use small area template convolution to do approximation. G_x and G_y need a template each, so there must be two templates combined into a gradient operator. The two 3×3 templates used by Sobel are showed as

-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

Figure 2: Sobel 3X3 Edge mask

Every point in the image should use these two kernels to do convolution. One of the two kernels has a maximum response to the vertical edge and the other has a maximum response to the level edge. The maximum value of the two convolutions is used as the output bit of the point, and the result is an image of edge amplitude.

C. Object Detection:

Object detection is the process of finding instances of real-world objects such as faces, bicycles, and buildings in images or videos. The salient object is detected with the help of Active contour algorithm.

An active contour or snake is a curve defined in an image that is allowed to change its location and shape until it best satisfies predefined conditions. It can be used to segment an object by letting it settle –much like a constricting snake– around the object boundary. Active contours are computer generated curves that move within the image to find object boundaries under the influence of internal and external forces.

This procedure is as follows:-

Step 1: Snake is placed near the contour of Region of Interest (ROI)

Step 2: During an iterative process due to various internal and external forces within the image, the Snake is attracted towards the target. These forces control the shape and location of the snake within the image.

Step 3:An energy function is constructed which consist of internal and external forces to measure the appropriateness of the Contour of ROI, Minimize the energy function (integral), which represents active contour's total energy, The internal forces are responsible for smoothness while the external forces guide the contours towards the contour of ROI.

The movement of the snake is modeled as an energy minimization process, where the total energy E to be minimized consists of three terms:

$$E = \int_0^1 E(C(s))ds = \int_0^1 (E_i(C(s)) + E_e(C(s)) + E_c(C(s)))ds \quad (6)$$

The term Ei is based on internal forces of the snake; it increases if the snake is stretched or bent. The term Ee is based on external forces; it decreases if the snake moves closer to a part of the image we wish it to move to. The term Ec is based on constraint force.

$$E_i = c_1 \left\| \frac{dC(s)}{ds} \right\|^2 + c_2 \left\| \frac{d^2C(s)}{ds^2} \right\|^2 \quad (7)$$

$$E_e = -c_3 \left\| \nabla f \right\|^2 \quad (8)$$

The idea behind active contours, or deformable models, for image segmentation is quite simple. The user specifies an initial guess for the contour, which is then moved by image driven forces to the boundaries of the desired objects. In such models, two types of forces are considered - the internal forces, defined within the curve, are designed to keep the model smooth during the deformation process, while the external forces, which are computed from the underlying image data, are defined to move the model toward an object boundary or other desired features within the image. One way of describing this curve is by using an explicit parametric form, which is the approach used in snakes. This causes difficulties when the curves have to undergo splitting or merging, during their evolution to the desired shape. To address this difficulty, the implicit active contour approach, instead of explicitly following the moving interface itself, takes the original interface and embeds it in higher dimensional scalar function, defined over the entire image.

Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process used to assign a label to every pixel in an image such that pixels with the same label share certain characteristics. Simply the segmentation subdivides an image into its constituent regions or objects and the subdivision is carried depends on the problem being solved. The Grabcut algorithm is used for segment the desired object from the detected region.

D. Salient Object Detection:

Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process used to assign a label to every pixel in an image such that pixels with the same label share certain characteristics. Simply the segmentation subdivides an image into its constituent regions or objects and the subdivision is carried depends on the problem being solved. The Grabcut algorithm is used for segment the desired object from the detected region.

Grabcut chooses segmentation by iteratively revising foreground and background pixel assignments. At the start, Grabcut initializes a foreground and background model using the bounding box. Any pixel outside the bounding box can be confidently assigned to the background. Pixels within the bounding box are initially assigned to foreground. Using these initial assignments, Grabcut then develops a Gaussian mixture model (GMM) for the foreground and background respectively. Each GMM is made up of multiple components, which represent clusters of similar pixels. Grabcut then defines a distance measure between each pixel and the foreground and background models, based on the component most similar to the pixel in each model. Grabcut also calculates the colour distance between each pixel and its neighbours. Using these two distance measures, Grabcut models the entire image as a directed, weighted graph, where each pixel has an edge to a foreground source node, a background sink node, and each of its eight neighbours. The distance measures (D) to the foreground and background models weight the edges from a pixel to the source and sink nodes, respectively. These edges represent the cost of assigning an individual pixel to the foreground or background.

$$U(\alpha, k, \theta, z) = \sum_n D(\alpha_n, k_n, \theta, z_n) \quad (9)$$

$$D(\alpha_n, k_n, \theta, z_n) = -\log \pi(\alpha_n, k_n) + \frac{1}{2} \log \det \sum(\alpha_n, k_n)$$

$$+ \frac{1}{2} [z_n - \mu(\alpha_n, k_n)]^T \sum(\alpha_n, k_n)^{-1} [z_n - \mu(\alpha_n, k_n)] \quad (10)$$

Edges from a pixel to its neighbors have weights that consider both the color distance between the two pixels and the landscape of the pixel neighborhood. These pair wise edges (V) represent the cost of assigning neighboring pixels to different segments and are engineered with such that neighboring pixels are only given opposite assignments when there is a significant change in color over multiple pixels.

$$E(\alpha, k, \theta, z) = U(\alpha, k, \theta, z) + V(\alpha, z) \quad (11)$$

$$\beta = \left(2 \left\langle (z_m - z_n)^2 \right\rangle \right)^{-1} \quad (12)$$

This helps encourage smooth segmentations - one outlier pixel in an otherwise homogenous segment will have the same assignment as its neighbors even if it is a very different color. Grabcut then calculates the minimum cut of the constructed graph to find the minimum-cost segmentation (E) and re-assigns pixels to the foreground and background accordingly. The entire process then repeats until convergence by relearning the GMM models and constructing another graph, etc

$$E(\alpha, k, \theta, z) = U(\alpha, k, \theta, z) + V(\alpha, z) \quad (13)$$

E. Foreground Extraction:

Foreground extraction aims to isolate foreground objects from potentially confusing background. In this paper, foreground extraction results can be obtained easily, in which the saliency segmentation results of an image are mapped onto the original image.

IV. SIMULATION RESULTS

The simulation studies involve the segmentation of salient region from underwater images shown in Figure 3. The proposed segmentation algorithm is implemented with MATLAB.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

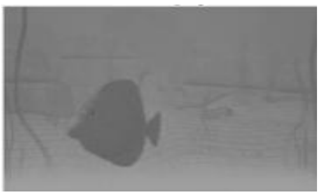

Figure 3.(a)

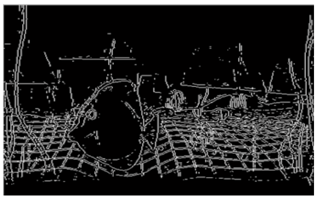

Figure 3.(b)

Figure 3.(c)

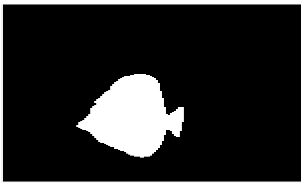

Figure 3.(d)

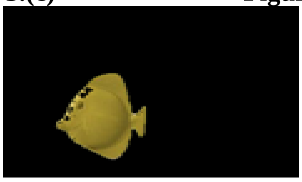

Figure 3.(e)

Figure 3: (a) Gray scale output of the given underwater image, (b) Edges are detected by Sobel Operator, (c) Salient Region detected by Active contour method, (d) Salient object segmentation by Grabcut method, (e) Foreground extraction of Salient Object

V. CONCLUSION AND FUTURE WORK

In this paper, a simple and efficient salient objects segmentation approach for underwater images by combining the conventional saliency detection approach and the Edge detection method is introduced. Unlike the conventional saliency detection methods, our proposed approach works well for underwater images. Even if numerous approaches for image segmentation are available, they are mainly limited to ordinary images and few approaches have been specifically developed for underwater images. This field of image processing is still a hot topic of research. The quality of images is directly affected by noise, pressure and temperature. This emphasizes the necessity of image segmentation. In the active contour algorithm, the main advantage is its simplicity, speed and allows high iterations which is done after the edges are detected by using the sobel edge detector. The result from the active contour algorithm is segmented by a technique called as Grabcut and the foreground extraction of the object from the image by simple mapping method. In the result, the foreground extraction of the desired object from an underwater image is done efficiently and quickly. In future, multiple objects can be detected, detect the things which are lost under the water by using advanced techniques and also track particular object underwater.

REFERENCES

1. Cheng, Ming, et al. "Global contrast based salient region detection." Pattern Analysis and Machine Intelligence, IEEE Transactions on 37.3 (2015): 569-582 .
2. M.Clerc, “The Swarm and the Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization,” In Proceedings of the IEEE Congress on Evolutionary Computation (CEC), pp. 1951-1957, 1999. (conference style)
3. Goferman, Stas, Lihi Zelnik-Manor, and Ayiellet Tal. "Context-aware saliency detection." Pattern Analysis and Machine Intelligence, IEEE Transactions on34.10 (2012): 1915-1926.
4. Rother, Carsten, Vladimir Kolmogorov, and Andrew Blake. "Grabcut: Interactive foreground extraction using iterated graph cuts."ACM Transactions on Graphics (TOG) 23.3 (2004): 309-314.
5. Chen, Tao, et al. "Sketch2Photo: internet image montage." ACM Transactions on Graphics (TOG) 28.5 (2009): 124
6. Avidan, Shai, and Ariel Shamir. "Seam carving for content-aware image resizing." ACM Transactions on graphics (TOG). Vol. 26. No. 3. ACM, 2007.
7. Christopoulos, Charilaos, Athanassios Skodras, and TouradjEbrahimi. "The JPEG2000 still image coding system: an overview. " Consumer Electronics, IEEE Transactions on 46.4 (2000): 1103-1127.
8. Achanta, Ravi, et al. "Frequency-tuned salient region detection." Computer vision and pattern recognition, 2009. cvpr 2009. IEEE conference on. IEEE, 2009.
9. Itti, Laurent, Christof Koch, and Ernst Niebur. "A model of saliency- based visual attention for rapid scene analysis." IEEE Transactions on Pattern Analysis & Machine Intelligence 11 (1998): 1254-1259.
10. Zhai, Yun, and Mubarak Shah. "Visual attention detection in video sequences using spatiotemporal cues."Proceedings of the 14th annual ACM international conference on Multimedia. ACM, 2006.
11. Hou, Xiaodi, and Liqing Zhang. "Saliency detection: A spectral residual approach." Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE, 2007.
12. Cheng, Ming-Ming, et al. "Efficient salient region detection with soft image abstraction." Proceedings of the IEEE International Conference on Computer Vision. 2013.
13. Caselles, f. Catté (catte), t. Coll, f. Dibos, "a geometric model for active contours in image processing", numer. Math., vol. 66, pp. 1-31, 1993.
14. Y. Hu, X. Xie, W.-Y. Ma, L.-T. Chia, and D. Rajan. Salient region detection using weighted feature maps based on the human visual attention model. Pacific Rim Conference on Multimedia, 2004.
15. Z. Wang and B. Li. A two-stage approach to saliency detection in images. IEEE Conference on Acoustics, Speech and Signal Processing, 2008.

AUTHORS PROFILE

Prasanna.S received her B.E. degree in Electronics and Communication Engineering from K.Ramakrishnan College of Engineering, Trichy, India in 2015. She is pursuing her M.E. degree in E.G.S Pillay Engineering College, Nagappattinam, India.

Durgadevi.S received her B.E. degree in Electronics and Communication Engineering from Trichy Engineering College Trichy, India in 2002 and M.E. degree in VLSI Design from Veltech Engineering College, Avadi, Chennai, India in 2009. She is currently working as Assistant professor in E.G.S Pillay Engineering College, Nagappattinam, India. She has published 6 articles in international and national journals and more than 22 papers in international and national conferences. He is a life member of the ISTE, a member of the IAENG.

Senthilkumar.S received his B.E. degree in Electronics and communication Engineering from Anjalai ammal mahalingam engineering college, Kovilveni in 2011 and M.E. degree in Nano Electronics from Sastra University, Tanjore, India. He is currently working as Assistant professor in E.G.S Pillay Engineering College, Nagappattinam, India. He has published 10 articles in international journals and more than 20 papers in international and national conferences. He is a life member of the ISTE, a member of the IAENG and a member of the IETE.